Separation and Classification of Heart–Lung Audio from Single-Channel Recordings

Filipi Kunz, Alexander Yoon, Austin Holt

Problem Statement

Chest recordings captured with a digital stethoscope often contain heart sounds and breath sounds at the same time, which makes it difficult to tell them apart and to recognize common abnormalities with confidence. This proposal aims to develop and evaluate a clear and compact pipeline that takes a single-channel recording, separates it into a heart track and a lung track, and assigns each track a broad label (for example, regular rhythm vs. irregular rhythm for the heart; normal breathing vs. wheeze or crackle for the lungs). The goal is to deliver a solution that separates each mixed recording into heart and lung tracks verified against the dataset's reference sources and further assigns broad labels with quantified uncertainty.

Data Source

The project will use the Heart and Lung Sounds Dataset Recorded from a Clinical Manikin using a Digital Stethoscope (HLS-CMDS), which provides 535 WAV files (15 s each at 22,050 Hz) recorded with a 3M Littmann CORE device on a CAE Juno manikin. The collection includes heart-only recordings, lung-only recordings, and mixed recordings; for each mixture, there are matched reference files for the heart and the lungs, along with simple metadata such as sound type, chest location, and simulated gender [1]. If preliminary studies indicate that HLS-CMDS does not provide sufficient sampling or class coverage for the planned analysis, the dataset described by Raza et al. [2] will be used as an alternative data source to classify different broad labels for the heart.

Methodology

Audio will be normalized and converted to spectrograms that expose patterns over time and frequency. Single-source sets will train compact templates for heart and lung energy. For each mixture, a soft mask over the spectrogram will assign time–frequency regions to the most likely source, then two waveforms will be reconstructed. Quality will be reported against provided references using waveform mean-squared error and spectrogram distortion, per file, and averaged. Separated tracks will be summarized into simple features and fed to a regularized logistic or tree model; model choice will use repeated cross-validation with bootstrap confidence intervals for accuracy and F1. A bandpass baseline will be implemented for both separation and direct classification to make gains attributable.

Expected Results

Separation is expected to reduce reconstruction error relative to the baseline and to improve source-specific clarity in the spectrograms, which should translate into higher downstream classification accuracy for both heart and lung labels. Results will be presented with accuracy, F1, and confusion matrices, together with confidence intervals from resampling, and brief notes on typical failure cases (e.g., very weak heart components or strongly overlapping frequency content).

References

[1] Torabi, Y., Shirani, S., and Reilly, J., Descriptor: Heart and Lung Sounds Dataset Recorded from a Clinical Manikin using Digital Stethoscope (HLS-CMDS). *IEEE Data Descriptions*, 2025.

[2] Raza, A., Mehmood, A., Ullah, S., Ahmad, M., Choi, G., and On, B., Heartbeat sound signal classification using deep learning. *IEEE Data Descriptions*, 2025.